UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified

1b. RESTRICTIVE MARKINGS: - -

AD-A224 535

DTIC
ELECTE
JUL 2 4 1990
D

3. DISTRIBUTION/AVAILABILITY OF REPORT: For Announcement Only

4. PERFORMING ORGANIZATION REPORT NUMBER(S): - -

5. MONITORING ORGANIZATION REPORT NUMBER(S): A 90-03

6a. NAME OF PERFORMING ORGANIZATION: Institute of Cognitive Science

6b. OFFICE SYMBOL (If applicable): - -

7a. NAME OF MONITORING ORGANIZATION: U.S. Army Research Institute for the Behavioral and Social Sciences

6c. ADDRESS (City, State, and ZIP Code): University of Colorado, Campus Box B-19, Boulder, CO 80309-0019

7b. ADDRESS (City, State, and ZIP Code): 5001 Eisenhower Avenue, Alexandria, VA 22333-5600

8a. NAME OF FUNDING/SPONSORING ORGANIZATION: see 7a.

8b. OFFICE SYMBOL (If applicable): PERI-BR

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: MDA903-86-K-0155

8c. ADDRESS (City, State, and ZIP Code): see 7b.

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
6.11.02B	74F	n/a	n/a

11. TITLE (Include Security Classification): Acquisition and Retention of a Letter Detection Skill

12. PERSONAL AUTHOR(S): Alice F. Healy, David W. Fendrich (U. of Colorado) Janet D. Proctor (Purdue University)

13a. TYPE OF REPORT: Announcement

13b. TIME COVERED: FROM 86/08 TO 90/07

14. DATE OF REPORT (Year, Month, Day): 1990, July

15. PAGE COUNT: 12

16. SUPPLEMENTARY NOTATION: Michael Kaplan, contracting officer's representative
Availability: Healy, A.F., Fendrich, D.W., and Proctor, J.D. (1990) Acquisition and (OVER)

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Mapping, Acquisition, Retention, Recognition

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This article reports on two experiments examining the acquisition and retention of a letter detection skill with a consistent mapping procedure. In the first experiment, subjects were trained (in from 0 to 4 sessions) to detect the letter "H" in displays containing random letters. Retesting was done after one month. Performance improved, and in some cases became more automatic, and the performance level was maintained over the retention interval. When tested with a prose passage, the high error rate on the word "the" was eliminated after training, and after the retention interval as well, regardless of the amount of training.

In the second experiment, two subjects were given 12 sessions of training, followed by a retention test six months later. One subject also received a retention test 15 months after acquisition. Performance improved dramatically with training, and substantial (but not quite complete) automaticity was achieved. Performance on the retention tests was close to (OVER)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION: Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL: Michael Kaplan

22b. TELEPHONE (Include Area Code): 202/274-8722

22c. OFFICE SYMBOL: PERI-BR

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE: UNCLASSIFIED

DTIC FILE COPY

90 07 23 091

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

16. Supplementary Notation (continued)

retention of a letter-detection skill. in Journal of Experimental Psychology, Learning, Memory, and Cognition. Volume 16, No. 2.

19. Abstract (continued)

the final acquisition level. The surprising lack of forgetfulness in the study was contrasted with the substantial forgetfulness typically found in studies of verbal learning.

Accesion For
NTIS CRA&I ✓
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution /
Availability Codes
Dist | Avail and / or Special
A-1 | 21

UNCLASSIFIED

 SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Journal of Experimental Psychology
Learning, Memory, and Cognition
1990, Vol. 16, No. 2, 270–281

Copyright 1990 by the American Psychological Association, Inc.
0278-7393/90/$00.75

Acquisition and Retention of a Letter-Detection Skill

Alice F. Healy and David W. Fendrich
University of Colorado

Janet D. Proctor
Purdue University

In two experiments, we examined the acquisition and retention of a letter-detection skill with a consistent-mapping procedure. In Experiment 1, subjects were trained from 0 to 4 sessions at detecting the letter H in displays containing random letters, and retesting occurred after a 1-month delay. Performance improved and in some cases became more automatic, and the performance level was maintained over the retention interval. When tested with a prose passage, the high error rate on the word THE was eliminated after training and after the retention interval, regardless of the amount of training. In Experiment 2, two subjects were given 12 sessions of training followed by a retention test 6 months later. For 1 subject there was also a retention test 15 months after acquisition. Performance improved dramatically with training, and substantial but not complete automaticity was achieved. Performance on the retention tests was close to the final acquisition level. The surprising lack of forgetting in this study was contrasted with the substantial forgetting typically found in studies of verbal learning.

In this investigation we are concerned with the acquisition and retention of a letter-detection skill. In previous research, letter-detection performance has been studied in two different contexts, one involving prose passages (e.g., Healy, 1976) and the other involving random letter displays (e.g., Schneider & Shiffrin, 1977). Although there has been a thorough investigation of the effects of training with random letters (including explorations of the development of automaticity), there has been essentially no research examining how training in that context affects subsequent performance in the prose context. Also, little is known about the durability of the effects of training in the letter-detection task (but see Rabbitt, Cumming, & Vyas, 1979). The present study examines the durability of the effects of training on letter detection, whether retention of the letter-detection skill depends on the amount of training or the achievement of automaticity, and the extent to which training in random letter displays influences detection in the prose context.

Dramatic forgetting is ubiquitous in verbal learning (see, e.g., Crowder, 1976), but forgetting seems to be considerably smaller in motor learning (see, e.g., McGeoch, 1942, and Naylor & Briggs, 1961, but also see, e.g., McGeoch & Melton, 1929) and relatively small in other studies of perceptual learning (e.g., Kolers, 1976). Perhaps, the learning resulting from detection training will be well retained, like motor and other perceptual learning. If so, the changes in detection performance resulting from detection training should be evident even after a relatively long delay without practice. On the other hand, if forgetting of the letter-detection skill is rapid, like verbal learning, the changes in detection performance may be transient and disappear after a delay.

Two types of letter processing have been distinguished in the literature: controlled processing, which requires attentional resources and cognitive effort, and automatic processing, which requires only minimal cognitive capacity and attention (e.g., LaBerge & Samuels, 1974; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Perhaps it will be necessary to train letter-level processing to the point of automaticity (so that letter information would be accessible without attentional resources) in order to attain superior long-term retention of the letter-detection skill. Alternatively, the amount of forgetting may not depend on whether automatic processing develops.

In previous studies of letter detection in prose, two striking findings have been well documented. First, a "word frequency disadvantage" has been found, in which letters occurring in

Portions of this article were read at the 28th Annual Meeting of the Psychonomic Society, Seattle, Washington, November 6, 1987.

This research was supported in part by National Science Foundation Grant BNS-8025020, United States Air Force Human Resources Laboratory Contract VF5744-022-001, and United States Army Research Institute Contract MDA903-86-K-0155 to the Institute of Cognitive Science at the University of Colorado, by Biomedical Research Support Grant RR07013-16 awarded to the University of Colorado by the Biomedical Research Support Grant Program (Division of Research Resources, National Institutes of Health), by a Council on Research and Creative Work Grant-in-Aid from the University of Colorado, and by National Institute of Mental Health Training Grant MH14617-08 to the University of Colorado. Alice F. Healy was supported by a Faculty Fellowship from the University of Colorado and a James McKeen Cattell Fund Sabbatical award during the preparation of this article.

We are indebted to Richard M. Shiffrin for help in designing a preliminary experiment and for many stimulating discussions about this research, to Ernest Mross for help with computer programming, to Antoinette Gesi for help with testing subjects and data analysis, to Lyle E. Bourne, Jr., for many valuable suggestions concerning these experiments, and to Henry L. Roediger III, Keith Rayner, Tom Carr, and three anonymous reviewers for helpful comments about this report.

Correspondence concerning this article should be addressed to Alice F. Healy, Department of Psychology, Campus Box 345, University of Colorado, Boulder, Colorado 80309.

JUN 27 '90 12:25 FROM CU BLDR - PSYCHOLOGY PAGE.003

very common words (such as THE) are missed more often than those occurring in less common words (such as THY) (e.g., Healy, 1976). Second, a "word inferiority effect" has been found, in which letters are more likely to be missed in correctly spelled words (again, such as THE) than in misspelled words (such as TEH) (e.g., Healy & Drewnowski, 1983). Underlying the explanations of these effects (see, e.g., Drewnowski & Healy, 1977; Healy & Drewnowski, 1983; Healy, Conboy, & Drewnowski, 1987; McClelland & Rumelhart, 1981) is the basic assumption that the failure to detect letters in common, correctly spelled words results from interactions of processing at the word and letter levels. This assumption not only predicts that enhancing processing at the word level may inhibit further processing at the letter level but also leads to the prediction that enhancing processing at the letter level will change the pattern of letter-detection errors. Indeed, in previous research with the letter-detection task (Healy, Oliver, & McNamara, 1987), the pattern of errors has been found to be changed as a result of practice. Specifically, a decrease in the overall error rate and in the size of both the word frequency disadvantage and the word inferiority effect was found as a function of repeated exposure to the same prose passage. However, this effect of practice might have been the result of familiarity with the specific passage rather than the result of improved letter-level processing, especially because the effects of passage familiarization have been found to be substantial in studies of proofreading for misspellings (see Levy, 1983; Levy & Begin, 1984; Levy, Newell, Snyder, & Timmins, 1986). Hence, it is important to construct a situation in which only letter processing is practiced, so that the effects of training at the letter level can be assessed.

The aim of the present study was to examine these issues concerning the acquisition and retention of a letter-detection skill by constructing a task analogous to that used with letter detection in prose but in which only letter processing was practiced. We achieved this end by developing a variant of the detection training paradigm developed by Schneider and Shiffrin (1977). Specifically, as in the prose letter-detection task, character sequences were rapidly presented on a computer terminal screen, and subjects pressed a response key when they detected the target letter (see, e.g., Healy, Oliver, & McNamara, 1987; Proctor & Healy, 1985). In this case, however, random letter sequences, rather than connected text, were used. Further, as in the detection training paradigm, frame size (the number of letters in each display) was varied, yielding slower and less accurate responding with larger frame sizes. Automatic processing was indexed by a decrease in the effect of frame size as a function of practice. Before and after detection training, subjects were exposed to the standard letter-detection task with a prose passage. We expected that the word frequency disadvantage and word inferiority effect would be evident before training. However, these effects should be reduced or eliminated after training, especially if subjects became automatic at letter detection. In our experiments subjects also returned for additional testing after a long delay interval. During this retention test they were given another exposure to the prose letter-detection task, as well as the random-letter task that they had practiced. We expected that the effects of training on letter-detection performance in both tasks would be well maintained across the delay interval if this skill resembles other perceptual and motor skills in its retention characteristics.

Experiment 1

In preliminary research (Healy, Fendrich, & Proctor, 1987) we found that the word frequency disadvantage in the letter-detection task with prose passages was large in a pretest but was eliminated on a posttest after detection training. One purpose of Experiment 1 was to compare the effect of two different amounts of detection training. Perhaps the word frequency disadvantage will be eliminated only if the subjects are given sufficient practice so that their performance at least approaches automaticity. To address this question, we included two experimental groups of subjects who were exposed to different amounts of detection training: either 2 or 4 days of training were administered before the posttest.

The second and most important purpose of Experiment 1 was to examine the permanence of the effects of detection training. Towards this end, we employed a retention-test phase approximately 1 month after the posttest. The retention tests included letter detection in a prose passage, followed by the detection training task. The retention test with the detection training task allowed us to examine the durability of the detection skill across a lengthy delay interval. The retention test with the prose passage allowed us to assess the durability of the changes in the pattern of letter-detection errors resulting from detection training. Perhaps the word frequency disadvantage will be eliminated at the end of training but will reappear in the retention test. Alternatively, the skill learned during detection training may be retained so well that the pattern of results on the posttest will persist to such an extent that the word frequency disadvantage will continue to be absent during the retention test.

Method

Subjects

Thirty-six students at the University of Colorado participated for course credit in Introductory Psychology and for payment at the rate of $5 per hour for any additional hours beyond the course requirement.

Stimuli and Apparatus

Detection training stimuli. The detection training displays were strings of 16 letters and two internal blank spaces (see Figure 1). This length corresponded to the approximate length of the letter-detection passage displays. Each string contained 2, 4, or 16 scrambled uppercase letters, depending on the frame size (2, 4, or 16), randomly interspersed with 14, 12, or 0 filler characters which were number signs (#). A target character (H) was present in half of the character strings. Two blank spaces were randomly placed in each string with the constraint that they could not occur in the first, the last or adjacent string positions. These blanks gave the displays the appearance of the three-word configuration of the letter-detection passages

FRAME SIZE	
2	## ###### #####HI#
4	##O ## H####I##M###
16	WYSEYIG PEO PCNUHE

Figure 1. Example frames in Experiment 1.

The nontarget letters (distractors) used were the same as those used in the second letter-detection passage except that they were scrambled at random.

Letter-detection passages. Three prose passages in uppercase type were employed. One passage was adapted from a passage of Winston Smith's novel, *The Stranger from the Sea.* The text contained 484 words, including 72 test words containing the target letter H. The word THE accounted for 36 of the test words. The remaining 36 test words were other lower frequency words containing a single H. At most, one test word occurred in each three-word display segment, and test words were located with equal frequency in all three positions.

Half of the test words of each type were misspelled, and two versions of the passage were produced by varying which half of the words were misspelled. As a result, a word–nonword comparison and an examination of the effects of word familiarity were made possible without the confounding variables of word length and frequency of occurrence in the text. Twelve nontarget, filler words also were misspelled so that incorrect spelling would not automatically signify the presence of a target. The same filler words were misspelled in both versions of the passage, according to a prescribed procedure. The last letter of the word was replaced with another letter, unless the last letter was a target. In that case, the first letter was replaced. Original letters and substitutes were paired so that the same substitution was always made for a given letter (e.g., THE was always misspelled THD), except when that substitution would produce another word. In those cases, an alternative substitute letter was selected.

The second passage was adapted from another portion of *The Stranger from the Sea.* It contained 783 words, including 48 occurrences of the word THE and 90 other lower frequency words containing a single H. Test words occurred with equal frequency in all three positions in a segment, and at most one test word appeared in each segment. The misspelling procedure described above was implemented, except that 15 filler words (rather than 12) were misspelled.

One of the first two passages was expanded and modified to create the third passage. The third passage contained 1,296 words, including 204 test words containing the letter H. The word THE accounted for 102 of the test words. The remaining 102 test words were other lower frequency words containing a single H. Half of the test words of each type (THE and other) and 24 additional filler words were misspelled by using the procedure described above. Test words of each type and spelling occurred equally often (n = 17) in each of the three word positions in a display segment of text.

Reading comprehension tests were constructed for each passage. Each test contained eight moderately difficult, four-alternative multiple choice questions.

Apparatus. Except for the reading comprehension questions which were presented on paper, all stimuli were presented on a Visual 200 cathode-ray tube (CRT) display screen linked to a PDP-11/03 computer system. The computer controlled stimulus presentation and recorded response latencies. Each subject responded by pressing a button held in his or her preferred hand. Measured from a viewing distance of 50 cm, the mean length of a line of text across all three passages was 4.96° of visual angle, and detection training stimuli subtended 5.27° of visual angle. Single uppercase letters subtended 0.23° horizontally and 0.46° vertically. A 0.34° space occurred between words and in detection training stimuli.

Procedure

General design. Three groups of 12 subjects each participated in two, three, or five sessions conducted over approximately 3 to 5 weeks. Group assignment was made according to a prescribed rotation based upon a subject's time of arrival. Because the experiment was conducted during two school terms (spring, summer), equal numbers of subjects from each term were assigned to each group in order to counterbalance any extraneous factors arising from different student populations. The three groups differed only in respect to the amount of detection training subjects received. The control group received no detection training, whereas the limited training group received 2 days of training (10 blocks) and the extensive training group received 4 days of training (24 blocks).

A standard sequence of tasks was used with all subjects, although the control group did not participate in the detection training phase. Table 1 shows the specific order and timing of tasks for each group. The experiment began with a pretraining prose passage letter-detection task. The first session continued with the initiation of the detection training phase. Because subjects in the control group received no detection training, they proceeded directly to the next task (posttest prose passage letter-detection task). Subjects in the limited training group performed five blocks of detection training during the first session, and an additional five blocks 2 days later in the second session. The extensive training group subjects received five blocks of training during both the first and fourth sessions, and seven blocks during both the second and third sessions. For the extensive training group, Sessions 1 and 2 were separated by 2 days, as were Sessions 3 and 4; 5 days separated Sessions 2 and 3.

After the detection training phase was completed, subjects immediately performed a posttraining letter-detection task with a second prose passage. A retention interval of 3 to 5 weeks then elapsed before subjects returned for the final (retention) phase of the experiment. At that time, a third passage was presented for a retention test of letter detection in prose. Next, subjects (including the control group) performed five blocks of the detection training task to evaluate retention of the letter-detection skill.

Detection training. A variation of the rapid serial visual presentation procedure was used to present letter strings briefly in the approximate center of the terminal display screen, as in the prose passage letter-detection task. Three frame sizes (2, 4, 16; number signs filled any remaining character spaces) were employed. Subjects were instructed to press a button as rapidly as possible whenever the target (H) occurred.

Table 1
Order and Timing of Tasks: Experiment 1

Group	Task: Pretest session	Training duration (sessions)	Posttest session	Retention duration (weeks)
Control	Session 1	0	Session 1	3–5
Limited	Session 1	2	Session 2	3–5
Extensive	Session 1	4	Session 4	3–5

Each training session was organized into several blocks of trials. A training block consisted of three sets of 52 trials (26 target, 26 nontarget), one for each frame size. Frame size order was random within each block. Stimulus exposure duration was 1,500 ms throughout training.

Prose passage letter-detection tasks. The text was presented three words at a time in the approximate center of the computer terminal screen by means of a variation of the rapid serial visual presentation procedure (Forster, 1970). Each three-word segment was presented for 1,500 ms. Within each training group, half of the subjects received one passage version, and half received the other. Subjects were instructed to read for comprehension and to press a button as rapidly as possible whenever the target letter occurred. Subjects searched for H in each passage. Passage order was counterbalanced across subjects. Reading comprehension questions were administered in a multiple-choice, paper-and-pencil format immediately following each passage.

Results

Scoring Procedures

Because the rapid serial visual presentation procedure has essentially no interstimulus interval, a delayed response to one stimulus can be registered during the presentation of the following stimulus. A response latency criterion was adopted to prevent including this type of response in the data. A response was considered to be correct (hit) and was included in the calculation of the response latency and accuracy data if the response was made during the presentation of a target stimulus and if the response latency exceeded 200 ms. Responses made after the first 200 ms of a display presentation that did not include a target letter were scored as false alarm errors. All responses with latencies under 200 ms were not scored and were eliminated from further analysis.

Detection Training

The proportion of hits and the median response latencies for hits were computed for each subject as a function of test block and frame size. Daily means for each training group are shown in Tables 2 and 3. The standard errors of the mean proportion of hits in Table 2 are .010 for the limited training acquisition sessions and .007 for the extensive training acquisition sessions. The standard errors of the mean response latencies in Table 3 are 10 ms for the limited training acquisition sessions and 8 ms for the extensive training acquisition sessions. The standard errors of the retention data from all three groups in Tables 2 and 3 are .015 and 13 ms, respectively. The proportion of false alarms was computed but not analyzed further because of their very low frequency (mean = .03 for the limited training acquisition sessions, .02 for the extensive training acquisition sessions, and .02 for the final retention session of all three groups). Because of problems of interpretation due to ceiling effects on accuracy with Frame Sizes 2 and 4, we present the statistical analyses of the response latency data only.

Table 2
Mean Proportion of Hits as a Function of Training Group, Frame Size, and Day of Training: Experiment 1

Group	Frame size	Day of training 1	2	3	4	Retention
Control						
	2					.98
	4					.93
	16					.77
Limited						
	2	.99	.98			.98
	4	.94	.96			.93
	16	.64	.73			.76
Extensive						
	2	.99	1.00	.99	1.00	.99
	4	.95	.98	.98	.98	.99
	16	.74	.82	.85	.89	.88

Table 3
Average Median Response Latency (in Milliseconds) as a Function of Training Group, Frame Size, and Day of Training: Experiment 1

Group	Frame size	Day of training 1	2	3	4	Retention
Control						
	2					676
	4					781
	16					1,003
Limited						
	2	660	622			624
	4	770	728			729
	16	1,017	992			992
Extensive						
	2	687	631	628	600	633
	4	805	721	711	685	718
	16	1,000	959	931	915	903

Training phase. Subjects received different amounts of detection task training, depending on their condition. As a result, an overall analysis of training including all subjects could not be performed. Instead, the data from the training period for the limited (2-day) and extensive (4-day) training groups were initially analyzed separately to evaluate the development of processing automaticity.

For both the limited and extensive training groups, the hit rate was highest with small frame sizes and increased as training progressed. The effect of frame size diminished with training, and thus some progression toward automaticity did occur. This shift toward automaticity was minimal, however; the magnitude of the frame size effect was reduced only slightly, and a substantial difference remained between Frame Size 4 and Frame Size 16 when training ended.

As opposed to the accuracy measure, the response latency data of the limited training group gave no evidence for improved automaticity in a 2 × 3 (Day of Training × Frame Size) analysis of variance. The main effects of day of training, $F(1, 11) = 10.31$, $p < .01$, and frame size, $F(2, 22) = 292.17$, $p < .001$, were significant, but the Day of Training × Frame Size interaction did not even approach significance.

A similar pattern was present for the extensive training group in a 4 × 3 (Day of Training × Frame Size) analysis of

27 '90 13:27 FROM CU BLDR - PSYCHOLOGY PAGE.005

variance. Response latencies decreased with training, $F(3, 33) = 17.44$, $p < .001$, and responses were slower to larger frame sizes, $F(2, 22) = 219.33$ $p < .001$, but the Day of Training × Frame Size interaction was not significant.

Retention phase. To evaluate the extent to which the effects of detection training were retained over time, the limited training and extensive training groups' performances from the last detection training session were compared with those from the retention session. These data are included in Tables 2 and 3. The previous separate analyses of detection training did not allow a direct comparison of the degree to which automaticity was attained by the two groups. The current retention analyses do provide this comparison and indicate that although full automaticity was not achieved, the extensive training group reached a significantly greater degree of automaticity than did the limited training group.

Subjects who received extensive training had higher hit rates and smaller frame size effects on hit rates (greater automaticity) than subjects who received limited training. There was no difference between the hit rates on the last training session and the retention test, reflecting both groups' almost complete retention of letter-detection skills in the detection training task over 3 to 5 weeks. Further, the level of automaticity did not change over the retention period for either group.

Response latencies for the last training session and the retention test were compared in a 2 × 2 × 3 (Training Group × Day × Frame Size) analysis of variance. This analysis yielded only two significant effects, the main effect of frame size, $F(2, 44) = 488.29$, $p < .001$, and the Training Group × Frame Size interaction, $F(2, 44) = 6.58$ $p < .01$. Response latencies increased as frame size increased, and this effect was larger for subjects given limited training than for those given extensive training, suggesting more automatic responding for the subjects exposed to more training. This interpretation is supported by a trend analysis that revealed a significant Training Group × linear Frame Size interaction component, $F(1, 22) = 8.11$, $p < .01$.

Although subjects in the control group received no detection training during the main training sessions, they did receive five blocks of training during the retention test session. These data were used in a final set of analyses to compare the control group's performance during one session of detection training with the performance after a retention period for groups that had received limited or extensive training. To provide a more detailed account of any changes in performance, the data from each block of trials were used, rather than the daily averages used previously. These data are shown in Tables 4 and 5. The standard error of the mean proportion of hits in Table 4 is .017 and of the mean response latencies in Table 5 is 19 ms, as determined by analyses of variance.

The hit rate of the control group improved over the session, whereas the performance of the limited training group decreased slightly. Further, the extensive training group had a higher hit rate in general and a smaller frame size effect than the control and limited training groups had. For the proportion of hits, therefore, the limited training group performed no better after the retention interval than the control group performed during its initial training, but the extensive training group performed at a higher level and was more automatic than the groups receiving less training.

The data for response latency from the overall 3 × 5 × 3 (Training Group × Block × Frame Size) analysis of variance were partitioned into comparisons of the control group versus the limited training group, and the combination of the control group and the limited training group versus the extensive training group. Only the Training Group × Block term in the comparison of the control and limited training groups was significant, $F(4, 132) = 4.13$, $p < .01$. As with the accuracy measure, the control group's performance improved across blocks, whereas the limited training group's performance worsened somewhat. When the control and limited training groups were combined and their response latencies were compared with those of the extensive training group, only the interaction of training group and frame size was significant, $F(2, 66) = 6.78$, $p < .01$. A trend analysis revealed that this interaction included a significant Training Group × Linear Frame Size component, $F(1, 33) = 8.58$, $p < .01$. The main effect of training group was not significant. Although the overall level of response latency did not differ between the extensive training group and the groups receiving less training.

Table 4
Mean Proportion of Hits in Retention Test as a Function of Training Group, Frame Size, and Trial Block: Experiment 1

Group	Frame size	Trial block 1	2	3	4	5
Control						
	2	.96	.98	.98	.97	.99
	4	.90	.91	.94	.93	.96
	16	.74	.76	.79	.76	[illegible]
Limited						
	2	.99	1.00	.97	.97	.97
	4	.97	.95	.95	.90	.90
	16	.76	.79	.76	.74	.74
Extensive						
	2	1.00	.99	1.00	1.00	.99
	4	1.00	.98	.99	.98	.98
	16	.93	.88	.88	.84	.88

Table 5
Average Median Response Latency (in Milliseconds) for Retention Test as a Function of Training Group, Frame Size, and Trial Block, Experiment 1

Group	Frame size	Trial block 1	2	3	4	5
Control						
	2	699	700	650	664	665
	4	864	791	756	746	745
	16	1,026	1,013	1,013	991	972
Limited						
	2	625	600	597	641	656
	4	703	721	733	755	733
	16	984	1,006	1,013	1,001	958
Extensive						
	2	610	634	633	633	656
	4	709	719	708	724	732
	16	918	924	916	860	897

the magnitude of the frame size effect was smaller for the extensive training group, and thus, automaticity of processing was greater.

Letter Detection in Prose

The proportion of targets detected (hits) was computed for each subject as a function of test-word type (THE/other) and test-word spelling (correct/misspelled). As in the detection training analyses, a response latency criterion of 200 ms was adopted. For the present analyses, all latencies under 200 ms were treated as failures to respond, as in previous studies. Group means are shown in Table 6. The standard error of the mean proportion of hits in Table 6 is .023, as determined by an analysis of variance. The mean proportion of false alarm errors overall was quite small (mean = .03); hence, these false alarm data will not be discussed further.

Because all groups received the same three passage tests, the data from all subjects could be combined in one overall 3 × 3 × 2 × 2 (Training Group × Test × Word Type × Spelling) analysis of variance.

The main effects of word type, $F(1, 33) = 4.43$, $p < .05$, and spelling, $F(1, 33) = 85.65$, $p < .001$, were significant, with a greater proportion of hits overall on THE relative to other, less common words (a word frequency advantage), and on misspelled words than on correctly spelled words (a word inferiority effect). Most important, the Word Type × Spelling interaction was significant, $F(1, 33) = 75.34$, $p < .001$. As in previous experiments, subjects made the lowest proportion of hits on correctly spelled instances of the common word THE, a somewhat greater proportion of hits on other correctly spelled words, and the greatest proportion of hits on misspelled instances of THE and other words. Thus, the word inferiority effect was greater for the word THE than for other words.

The results of the pretraining letter-detection task nicely match those of previous studies (Healy, Oliver, & McNamara, 1987; Proctor & Healy, 1985). Targets were less likely to be detected when the target was in the context of a correctly spelled, very high-frequency word, such as THE. This replication extends the generality of the word frequency disadvantage and word inferiority effect to new passages and a new target letter.

Table 6
Mean Proportion of Hits as a Function of Training Group, Test, Word Type, and Spelling: Experiment 1

Group	Word	Pretest Cor	Pretest Mis	Posttest Cor	Posttest Mis	Retention Cor	Retention Mis
Control							
	THE	.54	.81	.67	.92	.74	.91
	other	.67	.75	.68	.72	.70	.75
Limited							
	THE	.39	.76	.57	.92	.60	.88
	other	.63	.70	.68	.73	.62	.68
Extensive							
	THE	.47	.81	.70	.91	.70	.95
	other	.68	.73	.75	.79	.76	.81

Note. Cor = correct; Mis = misspelled.

Performance changed as a function of test. The main effect of test day was significant, $F(2, 66) = 18.49$, $p < .001$, with proportion of hits increasing from the pretest to the posttest and then remaining relatively unchanged in the retention test. Also, test interacted with word type, $F(2, 66) = 32.39$, $p < .001$, and with spelling, $F(2, 66) = 3.62$, $p < .05$. On the pretest, fewer targets were detected with the word THE than with other words (a word frequency disadvantage), but on the posttest and retention test the opposite pattern occurred (a word frequency advantage), due primarily to a large increase in accuracy with the word THE (correctly spelled and misspelled) but only a modest increase in accuracy with other words. This type of reversal did not occur for the effect of spelling, but the difference between the proportion of hits made on misspelled words and on correctly spelled words decreased across tests. That is, the word inferiority effect decreased in magnitude with subsequent tests, although it remained substantial.

The length of training on the detection task had no significant effect on the proportion of hits for letter detection in prose. Neither the main effect of training group nor any of its interactions were significant.

Discussion

As in our preliminary study (Healy, Fendrich, & Proctor, 1987), the word frequency disadvantage was eliminated after detection training, and the word inferiority effect was reduced in magnitude. However, Experiment 1 provides no support for the hypothesis that the change in these effects was due to the detection training itself. The group receiving the most extensive detection training did not perform differently on letter detection in prose than did the groups that received limited or no training. In fact, the earliest loss seemed to occur for the control group, which received no detection training prior to the presentation of the passages. This finding suggests that experience with letter detection in prose itself is the critical factor. Further, passage familiarity cannot be the basis for this effect because a given subject saw a different passage at each testing. Most crucially, it should be noted that the change in performance was not short lived: the word frequency disadvantage did not reappear even after a retention interval of a month.

The disappearance of the word-frequency disadvantage as a result of experience with the prose letter-detection task may at first appear to be problematic for the unitization hypotheses (see, e.g., Healy, Oliver, & McNamara, 1987) because these hypotheses were developed specifically to account for the preponderance of letter-detection errors on frequent words. However, in fact, the findings from Experiment 1, although unexpected, do not pose a serious threat to the unitization hypotheses. According to these hypotheses, text is processed in parallel at the level of letters and words. Because of their familiar visual configuration, very common words like THE may be identified before their component letters. Once a word unit has been identified, the processing of the component letter units is terminated even if they have not yet reached the point of identification. This premature termination leads to errors on the letter-detection task and is caused by the pull of the text resulting from the comprehension processes. In this way the unitization hypotheses can account for the preponderance of letter-detection errors on common words like THE, the word frequency disadvantage. How can these hypotheses accommodate the loss of the word frequency disadvantage found with prior exposure to the prose letter-detection task? An explanation of this finding can be made simply by proposing that the pull of the text can be weakened with practice at the prose task, so that subjects learn to continue processing at the letter level even when word-level processing has already been completed. In other words, the compulsion to move on in a text at the expense of letter-level processing apparently can be diminished with experience at a task requiring letter identification in the context of reading prose. Detection training outside the prose context presumably cannot affect the compulsion to move on because there is no text, and hence no pull of the text, in that situation.

Although extensive detection training did not have a greater effect on the prose task than did limited training, it did produce greater automaticity in Experiment 1. Full automaticity was not obtained, however. We were surprised by this finding because it has been said that automaticity frequently develops in about 200 consistent-mapping trials or after 2 hr of training (Schneider & Fisk, 1983). Our extensive training group had considerably more practice than required by these norms. Further, previous studies (e.g., Dumais, 1979; Schneider & Shiffrin, 1977, Experiment 2) obtained automaticity for response latency with processing loads (Memory Set Size × Frame Size) of up to 16 characters (and with frame sizes of up to 16 characters in the study by Dumais), and the amount of training in these studies was comparable to that used in Experiment 1. One difference between the present experiments' procedure and that of Schneider and Shiffrin (1977) is the physical arrangement of the stimuli. Whereas they used a central fixation and presented the letters in a square around fixation, in the present experiment we displayed the stimuli in a string extending from left to right, a format similar to that found in normal text. In addition, the display size used by Schneider and Shiffrin allowed subjects to view all characters with high acuity in a single fixation, whereas the display size used in the present experiment presumably required several fixations in order for all characters to be seen with high acuity (see Shiffrin & Schneider, 1977, p. 166, for a discussion of this issue).

We wondered whether the extent to which the stimulus falls in peripheral vision, the density of the letters, or some other aspect of the stimulus itself precluded the development of complete automaticity in Experiment 1. In order to test this hypothesis, we conducted a follow-up experiment (see Healy, Fendrich, & Proctor, 1987) which made use of distractor letters (O) that were maximally discriminable from the target letter (H). Specifically, this experiment included detection training like that used in Experiment 1 and the same procedure except that the distractor letters were always O. In the context of these distractor letters, unlike the random distractor letters used in Experiment 1, we predicted that the target letter would "pop out" (see, e.g., Gardner, 1973; Treisman & Patterson, 1984) and that no disadvantage for Frame Size 16 would be evident even with minimal training, unless the disadvantage was due solely to stimulus display characteristics which could not be overcome. In fact, we found that the large disadvantage for Frame Size 16 was eliminated in this follow-up experiment, so that performance was no worse (indeed was better) on Frame Size 16 than on Frame Size 4. Therefore, we concluded that the frame size effect, and hence the failure to find complete automaticity, in Experiment 1 cannot be attributed to artifacts concerning visual angle and other characteristics of the visual display.

In any event, the degree of automaticity did not seem to be related to the degree of long-term retention in Experiment 1. The limited training group, like the extensive training group, showed essentially perfect skill retention, even though there was evidence of automaticity (albeit weak evidence) only for the extensive training group. In fact, as mentioned above, all three groups of subjects, including the control group who was given no detection training, showed retention of the loss of the word frequency disadvantage over the 1-month retention interval.

Experiment 2

In Experiment 1 we found essentially no forgetting of the letter-detection skill in terms of both speed and accuracy over a 1-month retention interval. Hence, the major purpose of Experiment 2 was to assess retention of the detection skill over longer intervals.

In Experiment 1 we found only modest reductions in the frame size effects as a function of practice. However, subjects in Experiment 1 were given at most only 4 hr of practice. Therefore, a second purpose of Experiment 2 was to determine whether more intensive practice will lead to more dramatic changes in the frame size effects and, thus, to a greater degree of automaticity.

Two subjects were employed for Experiment 2. Each subject was given 12 1-hr sessions of detection training followed by a retention test 6 months after the training ended. One subject also received a second retention test 9 months after the first retention test (15 months after training ended). To verify further our findings from Experiment 1 with the prose passages, the subjects were also given pretest, posttest, and retention tests with the passage letter-detection task.

Method

Subjects

Two subjects were tested in this experiment. One subject (AG) was an undergraduate research assistant majoring in psychology at the University of Colorado. She had had extensive experience testing subjects in experiments on cognitive psychology, including concurrent participation as an experimenter in Experiment 1. Although generally familiar with the stimulus configurations and tasks because of her role as experimenter, she remained in a separate room from the subjects during most of the training and testing sessions and did not view the stimulus displays during the presentation of the stimuli to the subjects. The second subject (DS) had received his bachelor's degree from the University of Colorado within the previous year before training began. This subject had not been a psychology major and was not familiar with experimental psychology.

Design, Apparatus, and Procedure

All aspects of the apparatus and procedure are comparable to those for the analogous tasks of Experiment 1. In particular, the display duration was 1,500 ms in both the detection training and passage letter-detection segments of the experiment. As in Experiment 1, the subjects completed a comprehension test for each passage.

Testing was conducted in three phases. Phase 1, the acquisition phase, consisted of a pretest on letter detection in prose followed by 12 sessions of intensive training carried out by AG within a 28-day period and by DS within a 40-day period. Phases 2 and 3 consisted of retention tests. Phase 2 consisted of 1 day of testing 6 months after the last day of training of Phase 1. Phase 3, which applied only to AG, also consisted of a single day of testing 9 months after the Phase 2 test day.

Phase 1: Acquisition. The first day of Phase 1 included only a pretest with Version 1 of the first letter-detection passage used in Experiment 1. Each of the remaining 12 days of Phase 1 included seven blocks of detection training comparable to that employed in Experiment 1 (i.e., uppercase letters, with uppercase H the target). The final day of Phase 1 also included two posttest passage letter-detection tasks, following the usual seven blocks of detection training. The first posttest letter-detection passage was Version 1 of the second passage used in Experiment 1. The second posttest letter-detection passage was one version of the T-detection passage used in previous studies (see, e.g., Proctor & Healy, 1985). Again, this passage was converted to all uppercase letters, and the target was uppercase T. Unfortunately, the data from AG for this passage were lost because of a computer malfunction, so the data from this passage will not be reported for either subject.

Phases 2 and 3: Retention. Phases 2 and 3 each included a retention test for letter detection in prose, followed by seven blocks of the detection training task like that conducted in Phase 1. All stimuli were typed in uppercase, and uppercase H was the target in each task. Version 2 of the third passage used in Experiment 1 was presented in the prose letter-detection task of Phase 2. Version 2 of the pretest passage of Phase 1 was presented in Phase 3.

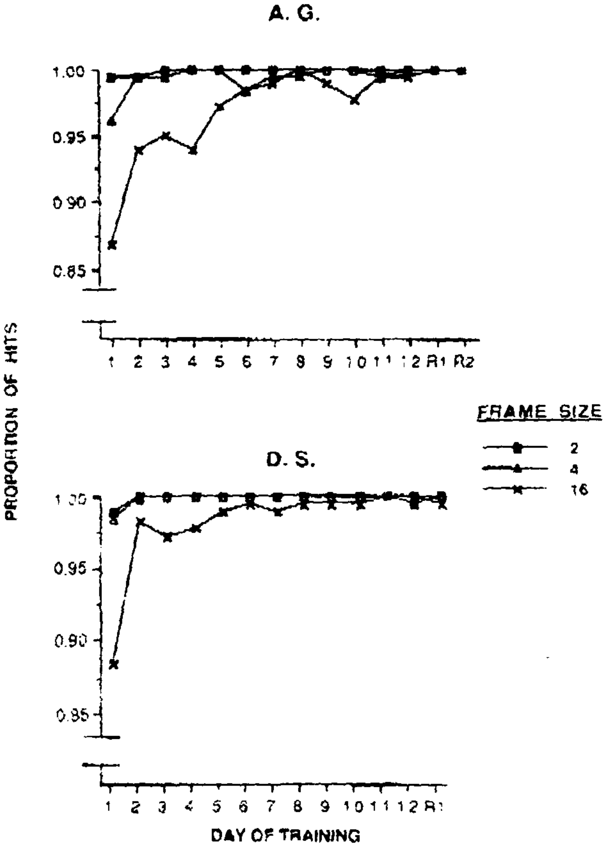

Figure 2. Mean proportion of hits as a function of day of training and frame size in Experiment 2: Data for AG in top panel, for DS in bottom panel. (R1 stands for the first retention test, and R2 for the second retention test.)

Results

Detection Training

Scoring procedures. The same scoring procedures were used as in Experiment 1. However, because only 2 subjects were tested in the present experiment, the factor of blocks (within days) rather than subjects, was treated as the random effect in two separate analyses of variance, one for each subject. Two types of analyses were conducted. The first type of analysis included data only from the first 12 sessions of training (Phase 1-Acquisition), with session and frame size as within-blocks factors. The second type of analysis included data only from the last session (Session 12) of acquisition training and the two retention sessions (Phases 2 and 3) for AG but only the one retention session for DS, again with session and frame size as within-blocks factors. Thus, in the first type of analysis there were 12 levels for the session factor, whereas in the second type of analysis there were only 3 levels for AG and 2 levels for DS. In both types of analyses there were three levels for the frame size factor (2, 4, and 16).

Accuracy data. Figure 2 shows the proportion of hits as a function of session and frame size for AG in the top panel and for DS in the bottom panel. The standard error of the mean proportion of hits in Figure 2 is .008 for AG and .005 for DS, as determined by analyses of variance. False alarms were also computed but were not analyzed further because of the low frequency of occurrence during acquisition (mean = .04 for AG and .02 for DS).

For both subjects, the hit rates for the three frame sizes are quite different initially but converge as the training progresses, so that by the final (12th) session, the hit rates are at the ceiling for all three frame sizes and stay at the ceiling during the retention tests.

Response latency data. Figure 3 presents the means of the median response latencies for the hits as a function of session and frame size, again for AG in the top panel and DS in the bottom panel. The standard error of the mean response latencies in Figure 3 is 20 ms for AG and 22 ms for DS, as determined by analyses of variance. In the analysis of variance for the acquisition period, there were significant main effects of session, $F(11, 66) = 37.93$, $p < .001$, for AG, and $F(11, 66) = 11.20$, $p < .001$, for DS, and frame size, $F(2, 12) = 204.92$, $p < .001$, for AG, and $F(2, 12) = 922.84$, $p < .001$, for DS, as well as a significant interaction between session

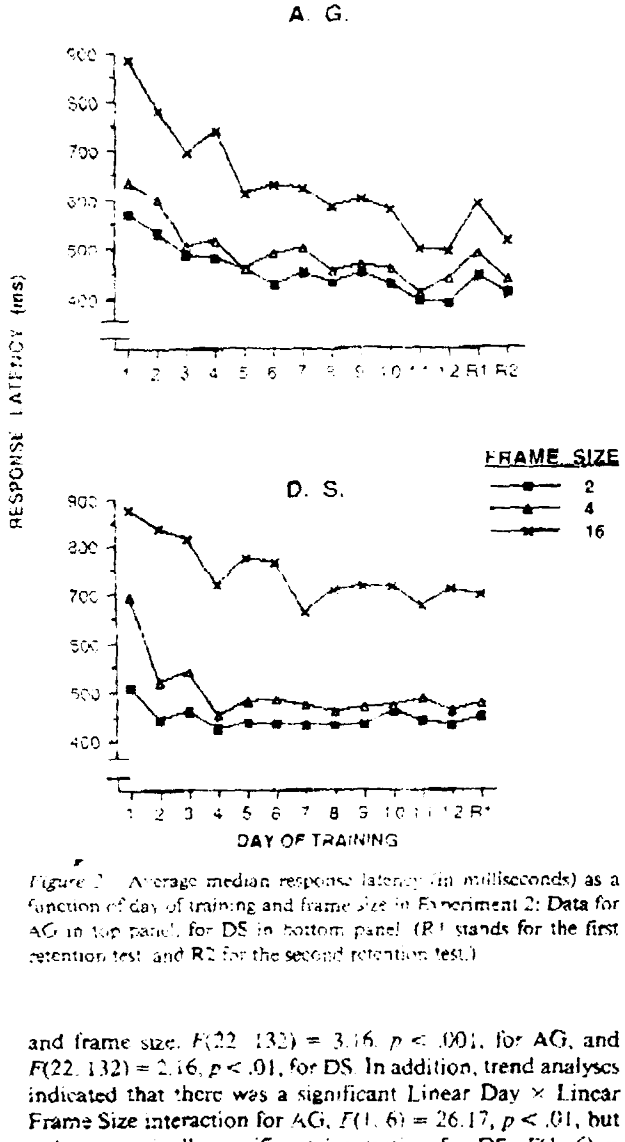

Figure 2. Average median response latency (in milliseconds) as a function of day of training and frame size in Experiment 2: Data for AG in top panel, for DS in bottom panel. (R1 stands for the first retention test and R2 for the second retention test.)

and frame size, $F(22, 132) = 3.16$, $p < .001$, for AG, and $F(22, 132) = 2.16$, $p < .01$, for DS. In addition, trend analyses indicated that there was a significant Linear Day × Linear Frame Size interaction for AG, $F(1, 6) = 26.17$, $p < .01$, but only a marginally significant interaction for DS, $F(1, 6) = 3.79$, $p < .10$. As for the hit rate, the frame size effect diminished as the sessions progressed, but in this case the effect was not eliminated entirely at the end of acquisition training, presumably because the latencies had not reached their lowest possible level. Note that all three frame sizes converged for AG, but only the smaller two frame sizes converged for DS. In any event, the significant interaction does support the hypothesis that a degree of automaticity was achieved, but the difference in frame sizes at the last session, especially for DS, suggests that automaticity was not complete.

For the second analysis, which compared the last session of the acquisition period to the retention session(s), there was a significant main effect of frame size, $F(2, 12) = 43.82$, $p < .001$, for AG, and $F(2, 12) = 96.15$, $p < .001$, for DS, with the smallest latencies for Frame Size 2 and the largest for Frame Size 16. In addition, for DS there was no effect of session, $F(1, 6) < 1$, but for AG there was a significant main effect of session, $F(2, 12) = 8.68$, $p < .01$, with shorter latencies for the final acquisition session and the final retention session than for the initial retention session. Planned comparisions for AG revealed no significant difference between the first (final acquisition) and third (final retention) tests, but the average of the first and third tests did differ from that of the second (initial retention) test, $F(1, 6) = 13.70$, $p = .01$. Thus, for AG the final retention test 15 months after acquisition yielded performance comparable to that at the end of training, suggesting essentially no forgetting, although there was a significant performance decrement after the first 6-month interval for that subject. Alternatively, after the first 6-month delay, there was significant forgetting evident at the retention test for AG, but that test provided a reminder which boosted performance back to the level attained at the final acquisition session. In contrast, no forgetting was evident at the 6-month retention test for DS.

Letter Detection in Prose

The scoring procedure for this task was the same as that used in Experiment 1. However, because this experiment included only 2 subjects, no statistical analysis was conducted.

The proportion of hits and false alarms was computed for each of the tests of letter detection in prose. The false alarm rate was small for each of the tests (mean = .03 for AG and mean = .02 for DS).

The proportion of hits as a function of test, word type, and spelling is shown in Table 7. The most striking aspect of these data is the increase in hit rate from the pretest to the posttest. This improved performance is maintained during the retention tests for AG but not fully maintained for DS. It is also noteworthy that the data from the pretest are consistent with both a word inferiority effect for THE (but not for other words) and a word frequency disadvantage. The data are generally inconsistent with a word frequency disadvantage and word inferiority effect for the posttest for both subjects and for the retention tests for AG.

Table 7
Proportion of Hits as a Function of Test, Word Type, and Spelling: Experiment 2

		Subject AG		Subject DS	
Test	Word	Cor	Mis	Cor	Mis
Pretest					
	THE	.72	.78	.61	.83
	other	.78	.72	.78	.83
Posttest H					
	THE	.96	.92	.92	1.00
	other	.96	.98	.96	.93
Retention 1					
	THE	.94	1.00	.80	1.00
	other	.86	.94	.94	.94
Retention 2					
	THE	1.00	1.00		
	other	1.00	1.00		

Note. Cor = correct; Mis = misspelled.

Thus, the results for these 2 subjects are generally consistent with those from the previous experiments in two respects. First, the effects of training are large and are retained throughout long periods of disuse (in this case up to 15 months). Second, the standard word frequency disadvantage is eliminated at the posttest for both subjects and remains absent at the subsequent retention tests for AG (but not DS).

Discussion

The 2 subjects employed in this experiment elucidated two important effects of training suggested, but not clearly demonstrated, in Experiment 1, which involved a greater number of subjects but substantially less training and a shorter retention interval. First, performance, both in terms of speed and accuracy, improved dramatically as training progressed. Although it appeared from Experiment 1 that performance might have reached an asymptotic level with as little as 4 days of training, it is clear instead that performance steadily increased throughout the 12-session training period. Although a frame size effect persisted for response latencies at the end of training, especially for DS, the effect was substantially reduced for latencies and was eliminated for errors. Hence, it is clear that the responses of AG and probably DS became more automatic as a result of the training. Second, and most interesting, the large improvements in performance were maintained with essentially no decline over a 6-month retention interval for DS and over a 15-month interval for AG, with only one refresher training session intervening between the training and AG's final retention test. This latter finding suggests that the perceptual skill of letter detection more closely resembles motor learning rather than verbal learning in its retention characteristics. The extremely large degree of retention evident here is surprising and certainly worth further exploration.

General Discussion

We can best summarize our findings by dividing them into three subsets: those concerning letter detection in prose, the role of automaticity, and long-term retention.

Letter Detection in Prose

In preliminary research (Healy, Fendrich, & Proctor, 1987), we found that the word frequency disadvantage was eliminated and the word inferiority effect was reduced after detection training. We replicated this result in the present study, but we also found the same change in the pattern of detection errors when subjects were given no detection training but instead merely performed a pretest with the prose task. Moreover, these results were uninfluenced by the degree of detection training. Hence, the changes we observed cannot be attributed to enhanced letter-level processing alone but also to exposure to a level of processing higher than the letter. These findings are consistent with those of Kolers and Magee (1978) showing only a moderate amount of transfer from naming inverted scrambled letters to reading inverted text. Although higher level processing is implicated, we can rule out passage familiarity as a factor, because, unlike previous studies (e.g., Healy, Oliver, & McNamara, 1987), the changes in the pattern of letter-detection errors occurred even when subjects were tested with a new passage never seen previously (cf. Levy, 1983; Levy & Begin, 1984; Levy et al., 1986). Hence, we propose that exposure to a pretest with the prose letter detection task enables subjects to change the focus of their attention from the word or phrase levels to the letter level and thus to detect target letters that would have otherwise been overlooked because they are in common words.

These findings are consistent with the basic assumption, discussed in the introduction, that failures to detect letters in common, correctly spelled words result from interactions of processing at the word and letter level. As outlined earlier, this assumption leads to the prediction that enhancing processing at the letter level will change the pattern of letter-detection errors. Our results imply that such a change takes place only when practice occurs in the context of real words so that subjects can learn to focus their attention on the letter level. More specifically, with respect to the unitization hypotheses (see, e.g., Healy, Oliver, & McNamara, 1987), our findings suggest that the pull of the text caused by the comprehension processes can be weakened by practicing letter detection in a prose context, so that subjects can learn to continue processing a given word at the letter level even when that word has already been identified.

The Role of Automaticity

Although subjects did not show evidence for fully automatic responding in Experiments 1 and 2, all subjects did show clear signs of improvement with training, and the subjects in Experiment 2, especially AG, showed dramatic improvements approaching automaticity.

Our finding that the effect of frame size persisted even after extensive practice is consistent with the findings from an experiment by Rabbitt et al. (1979). These investigators independently manipulated both target (or memory) set size and display (or frame) size in their Experiment 2. There were three different numbers of targets (two, four, or eight) and three different numbers of letters in the display (two, four, or nine). After extensive practice (60,000 trials across 25 days) on a consistent-mapping visual search task, they found that the effect of target set size was eliminated, but the effect of display size remained.

Most crucially, the durability of the detection skill does not seem to depend on the development of automatic processing. We found in Experiment 1 essentially perfect skill retention for subjects given both limited and extensive training, although there was no evidence of automaticity for the limited training group. Further, we found in the same experiment that both groups of subjects maintained the loss of the word frequency disadvantage over a 1-month retention interval.

Long-Term Retention

Our most interesting results concern the long-term retention of the letter-detection skill. Subjects showed essentially no forgetting of the skill that they had acquired even after relatively long retention intervals. This finding was most dramatic for the subjects of Experiment 2, who showed substantial improvements in performance as a result of training. DS showed no loss in the performance level achieved after a 6-month retention interval, and AG showed no loss after a 15-month retention interval with only a single refresher training session (i.e., 6-month retention test) intervening between initial acquisition and the final retention test. Not only was there little forgetting of the detection skill, but the large change in performance on the prose task (the elimination of the word frequency disadvantage) was generally (not for DS) maintained over a lengthy retention interval, even when the change was caused by an experience of a relatively short duration (i.e., reading a single pretest passage).

The previous study in the literature most closely related to our own is the one by Rabbitt et al. (1979). This study employed a visual search task like our own and similarly examined the training and subsequent retention of the search skill. In Experiment 1 of this study, 60 subjects were exposed to 3 days of training, with 1,000 trials per day (a total of 3,000 trials, similar to the 3,744 trials given to subjects in the extensive training group of our Experiment 1). They were then retested after retention intervals of 2, 4, or 6 weeks. For some subjects the retention test involved the same target and distractors as used in training, whereas for other subjects a transfer test involving new distractors was employed. Subjects showed improvements in response latencies as training progressed and no increase in response latencies after the 2- and 4-week intervals. There was a significant increase after the 6-week delay, although the latencies in that case were shorter than those at the start of practice. Hence, the results pointed to substantial degrees of skill retention up to 4 weeks, as we found in Experiment 1. Further, the results indicated significant forgetting after a 6-week delay, as we found in our Experiment 2 for AG, but not for DS, after a 6-month delay. The superior retention we found in our Experiment 2 may be due to the fact that the subjects in that study were exposed to considerably more extensive practice (13,104 trials). Also, our finding no loss in AG's performance after a 15-month interval suggests that a limited amount of refresher training can maintain the skill even if there is initially some forgetting. Thus, although our findings are not inconsistent with those from the earlier study, the work by Rabbitt et al. seems to have underestimated the remarkable durability of the perceptual skill.

The negligible amount of forgetting found in our study of perceptual learning contrasts with the substantial forgetting found in traditional studies of learning (e.g., consider the rapid forgetting of three letters over an 18-s retention interval found by Peterson & Peterson, 1959). There are at least three interrelated distinctions between the present task and the traditional tasks, and any one or any combination of these distinctions may be responsible for the different patterns of forgetting. First, the tasks differ along the dimension that we will refer to as *skill versus knowledge* (see Bourne, Ekstrand, & Dominowski, 1971); others have labeled this dimension *operational versus declarative knowledge* (knowing how vs. knowing that, Ryle, 1949), or *procedural versus declarative memory* (Anderson, 1983). The subjects in our study learned a skill, whereas the subjects in the more traditional studies of verbal learning acquired knowledge. Second, the tasks differ in terms of the memory systems distinguished by Tulving (1985). Subjects in our task were engaging the procedural memory system, whereas subjects in the traditional tasks were making use of episodic memory. Third, our task was an implicit memory test, as opposed to the explicit memory tests used in the traditional tasks (see, e.g., Graf & Schacter, 1985). Thus, the letter-detection task we studied was a skill involving an implicit test of the procedural memory system. Other examples of long-term retention with little forgetting have involved pursuit-rotor motor skills (e.g., Jahnke & Duncan, 1956), reading inverted text (e.g., Kolers, 1976), and the word-fragment completion test (e.g., Sloman, Hayman, Ohta, Law, & Tulving, 1988; Tulving, Schacter, & Stark, 1982). The pursuit rotor task seems to fall unambiguously in the domain of skill; the reading of inverted text seems to be a clear example of procedural memory; and the priming of word-fragment completion has been used as an implicit measure of memory.

All three of these distinctions point to the involvement of procedural memory as the crucial factor leading to stable memory representations. In agreement with the theoretical position put forth by Kolers and Roediger (1984), we propose that memory representations cannot be divorced from the procedures which were used to acquire them and that the durability of memory depends critically on the extent to which the learning procedures are reinstated at test. Implicit memory tasks like ours that require the direct storage and retrieval of procedures should, according to this argument, be acquired and maintained with much greater facility than explicit memory tasks that involve procedural memory more indirectly, such as those that have been categorized as involving knowledge or episodic memory. For example, in the standard list learning experiment, the memory coding procedures used by subjects to store the list are not easily retrieved or reinstated at the time of test, unless the subjects employ specific mnemonic procedures, such as the method of loci, the keyword method, or the chunking method learned by the expert SF (Ericsson & Chase, 1982). In contrast, the procedures used by our subjects during acquisition are easily reinstated during the retention test because the subjects are performing the same task (i.e., letter detection) in both instances. This characterization of memory is consistent with theories of transfer-appropriate processing (e.g., Bransford, Franks, Morris, & Stein, 1979) and encoding specificity (Tulving & Thompson, 1973), both of which postulate that memory performance will be best when the procedures required at the retention test match those employed during learning.

This emphasis on procedural memory not only provides an explanation for the substantial degree of retention we found of the detection skill in our study but also helps explain another puzzling observation we have made. We have found that the pattern of errors on the prose letter detection task is influenced greatly by a previous experience with detection in prose but not by experience with detection in scrambled letters. The lack of an influence in the latter case could be explained by proposing that subjects use qualitatively different procedures to detect letters in the two contexts.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Bourne, L. E., Jr., Ekstrand, B. R., & Dominowski, R. L. (1971). *The psychology of thinking*. Englewood Cliffs, NJ: Prentice-Hall.

Bransford, J. D., Franks, J. J., Morris, C. D., & Stein, B. S. (1979). Some general constraints on learning and memory research. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 331–354). Hillsdale, NJ: Erlbaum.

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

Drewnowski, A., & Healy, A. F. (1977). Detection errors on *the* and *and*: Evidence for reading units larger than the word. *Memory & Cognition, 5,* 636–647.

Dumais, S. T. (1979). *Perceptual learning and automatic detection: Processes and mechanisms*. Unpublished doctoral dissertation, Indiana University.

Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory. *American Scientist, 70,* 607–615.

Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Perception & Psychophysics, 8,* 215–221.

Gardner, G. T. (1973). Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology, 4,* 130–155.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 501–518.

Healy, A. F. (1976). Detection errors on the word *the*: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 235–242.

Healy, A. F., Conboy, G. L., & Drewnowski, A. (1987). Characterizing the processing units of reading: Effects of intra- and interword spaces in a letter detection task. In B. Britton & S. Glynn (Eds.), *Executive control processes in reading* (pp. 279–296). Hillsdale, NJ: Erlbaum.

Healy, A. F., & Drewnowski, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 413–426.

Healy, A. F., Fendrich, D. W., & Proctor, J. D. (1987, November). *The effects of training on letter detection*. Paper presented at the 28th Annual Meeting of the Psychonomic Society, Seattle, Washington.

Healy, A. F., Oliver, W. L., & McNamara, T. P. (1987). Detecting letters in continuous text: Effects of display size. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 279–290.

Jahnke, J. C., & Duncan, C. P. (1956). Reminiscence and forgetting in motor learning after extended rest intervals. *Journal of Experimental Psychology, 52,* 273–282.

Kolers, P. A. (1976). Reading a year later. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 554–565.

Kolers, P. A., & Magee, L. E. (1978). Specificity of pattern-analyzing skills in reading. *Canadian Journal of Psychology, 32,* 43–51.

Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior, 23,* 425–449.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293–323.

Levy, B. A. (1983). Proofreading familiar text: Constraints on visual processing. *Memory & Cognition, 11,* 1–12.

Levy, B. A., & Begin, J. (1984). Proofreading familiar text: Allocating resources to perceptual and conceptual processes. *Memory & Cognition, 12,* 621–632.

Levy, B. A., Newell, S., Snyder, J., & Timmins, K. (1986). Processing changes across reading encounters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 467–478.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

McGeoch, J. A. (1942). *The psychology of human learning: An introduction*. New York: Longmans, Green.

McGeoch, J. A., & Melton, A. W. (1929). The comparative retention values of maze habits and of nonsense syllables. *Journal of Experimental Psychology, 12,* 392–414.

Naylor, J. C., & Briggs, G. E. (1961). *Long-term retention of learned skills: A review of the literature* (ASD-TR-61-390). Wright-Patterson AFB, OH: Advanced Systems Division.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

Proctor, J. D., & Healy, A. F. (1985). A secondary-task analysis of a word familiarity effect. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 286–303.

Rabbitt, P., Cumming, G., & Vyas, S. (1979). Improvement, learning and retention of skill at visual search. *Quarterly Journal of Experimental Psychology, 31,* 441–459.

Ryle, G. (1949). *The concept of mind*. London: Hutchinson.

Schneider, W., & Fisk, A. D. (1983). Attention theory and mechanisms for skilled performance. In R. A. Magill (Ed.), *Memory and control of action* (pp. 119–143). New York: North-Holland.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Sloman, S. A., Hayman, C. A. G., Ohta, N., Law, J., & Tulving, E. (1988). Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 223–239.

Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 12–31.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40,* 385–398.

Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 336–342.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Received January 30, 1989
Revision received April 13, 1989
Accepted July 3, 1989 ■

PEER REVIEW
(Subject matter expert review)

Date Due: ______________
Date Received: ______________
Date Returned: ______________

Reviewer (full name) Michael Kaplan

Organization Basic Research Office

Manuscript Title Acquisition and Retention of a Letter Detection Skill

Author or COR Alice F. Healy

I. RATINGS (extent to which criteria are met):

A. FOREWORD, EXECUTIVE SUMMARY (BRIEF): Are they clear? Are they consistent with the contents? Are they directed to target readers or users? Do they concisely highlight the important findings? N/A

No __ Moderately __ Yes __ Substantially __

B. INTRODUCTION, BACKGROUND, OBJECTIVES: Is the literature review relevant to the research conducted (necessary and sufficient)? Is the statement of the problem pertinent to the target audience? Is it clear? Is it supported by concrete data or evidence?

No __ Moderately __ Yes ✓ Substantially __

C. APPROACH, METHOD: Is it appropriate? Was there a valid experimental design and data collection plan? Are they competently described and were they competently executed?

No __ Moderately __ Yes ✓ Substantially __

D. RESULTS: Are they clearly presented? Is there appropriate use of tables and figures? Were proper statistics used? Were they used correctly?

No __ Moderately __ Yes ✓ Substantially __

E. DISCUSSION AND CONCLUSIONS: Do they follow from the data and literature review? Are they comprehensible? Are conclusions and recommendations warranted and usable by intended audience?

No __ Moderately __ Yes ✓ Substantially __

ARI Form 185, 20 Oct 88

Peer Review (cont)

II. Recommendations

___ Should not be published.

___ Return for reconsideration (e.g., reanalysis, additional data collection, or rewrite). Should not be published as is. (Comments may be made in Section III or as an enclosure.)

✓ Publish after minor revisions. (Comments may be made in Section III or as an enclosure.)

___ Publish as is.

My name may / may not appear on the inside cover as reviewer.

Reviewer's Signature

III. Comments on any of the above ratings or recommendations.